

# Die Beeinflussung in der Evaluationstätigkeit in der Schweiz und was die SEVAL dagegen tun kann

**Lyn Pleger und Fritz Sager** | Eine beträchtliche Anzahl Evaluierender in der Schweiz gab jüngst in einer Befragung an, bereits mit dem Versuch der Beeinflussung durch Auftraggeber oder andere Stakeholder konfrontiert worden zu sein. Dabei kommt dem Begriff der Beeinflussung, für den bislang nur diffuse Definitionen existieren, eine zentrale Bedeutung zu. Ziel dieses Beitrags ist es deshalb, den Begriff der Beeinflussung im Hinblick auf die Unabhängigkeit von Evaluationen zu präzisieren. Hierfür erfolgt zum einen eine Auswertung offener Antworten auf die Frage, was Schweizer Evaluierende unter Beeinflussung verstehen. Zum anderen wird das DRSB-Beeinflussungsmodell vorgestellt, das mittels eines heuristischen Ansatzes versucht, die durch die Evaluierenden erlebten Beeinflussungsformen voneinander abzugrenzen. Zudem wird aufgezeigt, was für Massnahmen zur Vermeidung von Beeinflussung aus der Sicht von Evaluierenden in Frage kommen und welche Rolle die SEVAL diesbezüglich einnehmen könnte. Die qualitative Auswertung zeigt, dass Beeinflussung oftmals subtil geschieht, dass die Abgrenzung zwischen positiver und negativer Beeinflussung in der Praxis schwierig ist und dass Schweizer Evaluierende eine mitunter sehr ähnliche Vorstellung davon haben, welche Funktionen die SEVAL bei Vorfällen von Beeinflussung erfüllen sollte.

## Inhaltsübersicht

- 1 Einleitung
- 2 Beeinflussung von Evaluierenden in der Schweiz
- 3 Ein heuristisches Modell der Beeinflussung
- 4 Präventive Massnahmen
- 5 Rolle der SEVAL und der Evaluationsstandards
- 6 Diskussion

## 1 Einleitung

Jeder zweite Evaluierende in der Schweiz war bereits mindestens einmal der versuchten Einflussnahme durch den Auftraggeber oder einen anderen Stakeholder ausgesetzt (Pleger/Sager 2016). Dieser Befund steht im Einklang mit den Ergebnissen anderer Studien aus verschiedenen Ländern, die ebenfalls nachweisen, dass viele Evaluierende oftmals mit Beeinflussungsversuchen konfrontiert sind (Morris/Clark 2013; Stockmann et al. 2011; The LSE GV314 Group et al. 2014). Diese Ergebnisse sind alarmierend, da sie das Fundament empirischer Forschung, die Unabhängigkeit der Forschenden und somit die Gütekriterien, auf denen empirische Forschung fusst, bedrohen. So kann die Einflussnahme von Stakeholdern auf den Evaluationsprozess die Einhaltung von Gütekriterien quantitativer und qualitativer Forschung verletzen. Die *Objektivität* als eines der drei zentralen quantitativen Gütekriterien (Raithel 2006), die in den qualitativen Methoden als Neutralität bezeichnet werden kann (Lincoln/Guba 1985; Pracke/Wurster 1999, 185), ist dabei besonders gefährdet.

Die Forderung nach Neutralität bedeutet eine Ausrichtung an der Frage «[w]ie [...] man den Grad erreichen [kann], zu dem die Ergebnisse der Forschung durch die Subjekte und Umstände der Forschung bestimmt sind und nicht durch Voreingenommenheiten, Motivationen, Interessen oder Perspektiven der Forschenden» (Lincoln/Guba 1985, 290; Prakke/Wurster 1999, 184). Aufbauend auf diesem Zitat liesse sich im Kontext der Evaluationsdurchführung *Unabhängigkeit* als ein der Neutralität äquivalentes Gütekriterium für die Evaluationstätigkeit definieren: Gemeint ist damit das Streben nach Forschungsergebnissen, *unabhängig von* Voreingenommenheiten, Motivationen, Interessen oder Perspektiven *aller in den Evaluationsprozess involvierten Akteure*. Wenn andere in den Evaluationsprozess involvierte Akteure versuchen, aktiv in den Evaluationsprozess einzugreifen, ist diese Unabhängigkeit nicht mehr gewährleistet und es findet eine Beeinflussung des Evaluierenden statt.

Ob und in welchem Ausmass sich Evaluierende in der Schweiz mit einer Beeinflussung durch in den Forschungsprozess involvierte Personen konfrontiert sehen, wurde in einer Umfrage unter den Mitgliedern der Schweizerischen Evaluationsgesellschaft (SEVAL) erhoben (Pleger/Sager 2016). Die Ergebnisse zeigen, dass viele Evaluierende bereits Beeinflussungsversuchen durch Auftraggeber und andere Stakeholder ausgesetzt waren, wobei diese Beeinflussung in verschiedenen Formen und Phasen des Evaluationsprozesses stattgefunden haben. Der verwendete Fragebogen enthielt neben geschlossenen Fragen auch zahlreiche offene Fragen, die durch die Evaluierenden sehr ausführlich und zahlreich beantwortet wurden. Die Antworten beinhalteten zudem eine Vielzahl innovativer Anregungen und verschiedene Wahrnehmungen. Die grosse Anzahl qualitativer Rückmeldungen und deren Ausführlichkeit sind Indikatoren für die Bedeutung der Thematik für die Evaluierenden, und die hohe Qualität der Rückmeldungen ermöglicht deren Auswertung. Viele offene Rückmeldungen bezogen sich auf den Begriff der Beeinflussung. Sie zeigen, dass unter Schweizer Evaluierenden kein einheitliches Verständnis dieses Begriffs vorherrscht. Die begriffliche Abgrenzung ist schwierig, da der Begriff «von seiner Semantik her nicht nur eine negative, sondern auch eine positive Bedeutung haben» kann (Stockmann et al. 2011, 49).

Das Anliegen dieses Artikels ist es deshalb, die Analyse der quantitativen Dokumentation von Beeinflussung um die subjektive Perspektive der Evaluierenden zu erweitern. Hierbei sind insbesondere die Frage nach dem Umgang mit Beeinflussung, die Frage nach möglichen präventiven Massnahmen zur Verhinderung von Beeinflussung und schliesslich die Klärung des Begriffs der Beeinflussung selbst, insbesondere im negativen Sinne, von Bedeutung.

Die Frage, wie sich Beeinflussung aus Sicht der Evaluierenden manifestiert, soll auch deshalb im Zentrum stehen, damit die nach wie vor diffuse Definition des Begriffs der Beeinflussung in diesem Zusammenhang präzisiert werden kann.

In praktischer Hinsicht versucht dieser Beitrag hinsichtlich potenzieller präventiver Massnahmen die Frage zu beantworten, welche Rolle die SEVAL aus Sicht der Evaluierenden in Bezug auf das Thema der Beeinflussung einnehmen sollte.

Nachfolgend werden im zweiten Kapitel zunächst die wichtigsten Ergebnisse der Studie zur Beeinflussung von Schweizer Evaluierenden (Pleger/Sager 2016) zusammengefasst. Als Basis für eine konzeptuelle Annäherung an den Begriff der Beeinflussung liegt der Fokus der Ergebnisse auf verschiedenen von Schweizer Evaluierenden erlebten Beeinflussungsformen. Im Anschluss daran erfolgt im dritten Kapitel eine theoretische Erläuterung des heuristischen DRSB-Beeinflussungsmodells. Darauf aufbauend werden ebenfalls im dritten Kapitel die von Schweizer Evaluierenden erlebten Beeinflussungsformen auf das DRSB-Modell übertragen. Im vierten Kapitel werden potenzielle präventive Massnahmen diskutiert. Darauf aufbauend wird im fünften Kapitel die aus Sicht der Evaluierenden wahrgenommene Rolle der SEVAL im Zusammenhang mit Beeinflussung dargestellt. Der Artikel schliesst mit einer Diskussion, in der weitere Anregungen der Evaluierenden im Zusammenhang mit Beeinflussung aufgenommen werden.

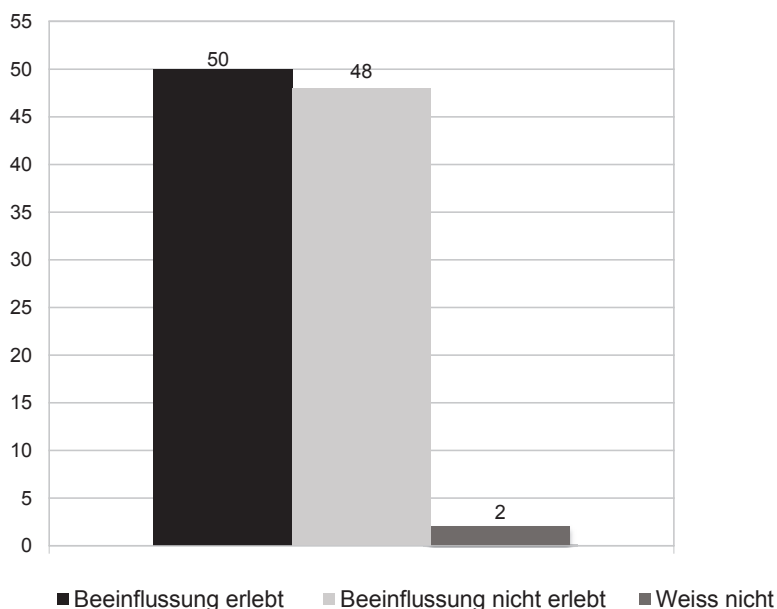
## **2 Beeinflussung von Evaluierenden in der Schweiz**

Die Befragung wurde zwischen dem 6. Mai und dem 6. Juni 2014 unter 473 SEVAL-Mitgliedern durchgeführt. Die Ausschöpfungsquote aller Angeschriebenen lag bei 33 Prozent (N=157). Mit Hilfe einer Filterfrage wurde sichergestellt, dass nur Auftragnehmer befragt wurden. Insgesamt gaben 22 Prozent (N=104) der Befragten an, als Auftraggeber tätig zu sein, was dazu führte, dass die Umfrage automatisch beendet wurde. Die geringe Anzahl Auftraggeber ist auch darauf zurückzuführen, dass bereits in der E-Mail mit dem Link zur Befragung ausdrücklich darauf hingewiesen wurde, dass sich die Umfrage ausschliesslich an Auftragnehmer richte und Antworten von Auftraggebern deswegen nicht berücksichtigt würden. Eine zweite Filterfrage stellte die Frage nach bereits erlebter Beeinflussung dar. Bei Evaluierenden, die angaben, noch nie beeinflusst worden zu sein, wurde die Umfrage beendet (siehe Pleger/Sager 2016 für eine ausführlichere Darstellung des Vorgehens). Inhaltlich orientierte sich der Fragebogen an vorgängig durchgeführten Befragungen aus den USA (Morris & Clark 2013), Grossbritannien (The LSE GV314 Group 2013) und Deutschland (Stockmann et al. 2011). In Anlehnung an den Fragebogen von Stockmann et al. (2011) wurde im Fragebogen der Begriff der Beeinflussung einleitend folgendermassen definiert:

Der Begriff der «Beeinflussung» kann für viele Menschen eine positive wie auch eine negative Bedeutung haben. Deshalb möchten wir Sie an dieser Stelle darauf hinweisen, dass unsere Fragen immer auf Beeinflussungen im *negativen* Sinne abzielen und Beeinflussung in dieser Umfrage ein Synonym für «Druck ausüben» darstellt (vgl. Pleger/Sager 2016).

Wie aus Abbildung 1 hervorgeht, gaben 50 Prozent der Evaluierenden (61 von 123) an, während ihrer Tätigkeit bereits mit Beeinflussung konfrontiert worden zu sein (Pleger/Sager 2016). Die Hälfte aller Evaluierenden erlebte demnach bereits Beeinflussungsversuche durch Stakeholder, am häufigsten durch Auftraggeber. Hinsichtlich der Beeinflussungshäufigkeit handelte es sich bei 90 Prozent der Befragten um keinen Einzelfall. Rund ein Drittel (32 %) der Evaluierenden gab an, während der Arbeit zwischen drei- bis sechsmal beeinflusst worden zu sein und 35 Prozent waren zweimal mit dem Versuch der Beeinflussung konfrontiert (Pleger/Sager 2016).

*Abbildung 1: Anteil Evaluierender, die bereits beeinflusst oder unter Druck gesetzt wurden (N=123)*



*Anmerkung:* Angaben in Prozent (in Anlehnung an Pleger/Sager 2016)

Neben der Beeinflussungshäufigkeit wurde im Fragebogen auch nach verschiedenen Beeinflussungsformen gefragt. Abbildung 2 fasst die vorgegebenen Antwortkategorien möglicher Formen der Beeinflussung durch den Auftraggeber sowie die dazugehörige Zustimmung externer und interner Evaluierender zusammen. In der Abbildung angegeben sind jeweils der Zustimmungsmittelwert (MW), dessen Standardabweichung (SD) sowie die Anzahl Antworten insgesamt auf diese Frage (N) (Pleger/Sager 2016). Hierbei wurden die Ergebnisse zudem separat nach internen und externen Evaluierenden dargestellt. Als interne Evaluierende in Abgrenzung zu externen Evaluierenden werden Evaluierende

bezeichnet, die bei der Institution angestellt sind, deren Tätigkeit sie evaluieren (Morris und Clark 2013). Es lässt sich deshalb annehmen, dass interne Evaluierende aufgrund ihrer persönlichen Verbindung zum zu evaluierenden Gegenstand anderen Druckformen ausgesetzt sind als externe Evaluierende. Aufgrund des ungleichen Verhältnisses der Stichprobengrösse zwischen externen Evaluierenden (84 %, N=104) und internen Evaluierenden (16 %, N=20) sowie aufgrund der insgesamt geringen Stichprobengrösse interner Evaluierender wird nachfolgend in der Diskussion keine weitere Unterscheidung zwischen externen und internen Evaluierenden vorgenommen.

Die Ergebnisse zeigen, dass die häufigste durch Schweizer Evaluierende erlebte Beeinflussungsform in der Umformulierung einzelner wertender Sätze im Evaluationsbericht besteht, die durch den Auftraggeber verlangt wurde. Am seltensten erlebten Evaluierende, dass der Auftraggeber verlangte, die Evaluationsergebnisse negativer darzustellen.

*Abbildung 2: Beeinflussungsformen der Auftraggeber*

Zustimmung zu möglichen Beeinflussungsformen des Auftraggebers (AG) <sup>a</sup>	Externe Evaluierende			Interne Evaluierende		
	MW	SD	N	MW	SD	N
AG hat verlangt, dass einzelne wertende Sätze im Evaluationsbericht umformuliert werden sollten.	3,0	0,9	50	2,8	0,9	8
AG hat verlangt, dass ich die Evaluationsergebnisse positiver darstelle.	2,4	0,8	52	2,1	1,1	8
AG verlangte ein anderes methodisches Vorgehen.	2,0	0,8	50	1,8	1,2	8
AG hat verlangt, dass andere Schlussfolgerungen aus den Ergebnissen gezogen werden sollten.	2,0	0,9	51	1,7	1,0	7
AG hat verlangt, dass ich Änderungen vornehme, die zu inhaltlichen Verzerrungen geführt hätten.	1,8	0,8	49	2,1	0,8	8
AG verlangte andere Datenquellen.	1,7	0,8	50	1,4	0,8	7
AG teilte mir im Vorhinein der Evaluation mit, welche Ergebnisse er erwartet.	1,6	0,8	50	1,9	1,1	8
AG hat verlangt, dass ich die Evaluationsergebnisse negativer darstelle.	1,4	0,7	51	1,1	0,4	8

<sup>a</sup> Angaben als Zustimmung auf einer Skala von 1 bis 4: 1 = Habe ich noch nie erlebt, 2 = Ja, habe ich einmal erlebt, 3 = Ja, habe ich mehr als einmal erlebt, 4 = Ja, habe ich oft erlebt.

Quelle: Pleger/Sager (2016)

Im Fragebogen wurde explizit darauf hingewiesen, dass Beeinflussung lediglich die negativ konnotierte Beeinflussung meint (vgl. oben). Trotz dieser Präzisierung wurde insbesondere durch die offenen Rückmeldungen der Evaluierenden die Schwierigkeit, Beeinflussung zu operationalisieren, deutlich. So zeigten die offenen Rückmeldungen, dass zunächst eine strikte Trennung zwischen positiver und negativer Beeinflussung erforderlich ist, um die Thematik der Unabhängigkeit von Evaluationen im Zusammenhang mit Beeinflussung untersuchen zu können (z. B. «Ich halte es für legitim, dass der Auftraggeber Korrekturen im Bericht vorschlägt.»). Ein Evaluierender fasste diese Abgrenzung folgendermassen zusammen: «Wie gesagt, hat Beeinflussung zwei Gesichter. Wenn Evaluationen einen Lernprozess der Beteiligten anregen sollen, müssen sie sich auch an den Aufnahmekapazitäten und Lernmöglichkeiten der Beteiligten orientieren und sich diesen bis zu einem gewissen Grad anpassen. Die Frage nach der Unabhängigkeit sollte diesen Balanceakt diskutieren». Ein anderer Evaluierender sagte, dass das dem Fragebogen zugrunde liegende Verständnis von Beeinflussung ein «ungenaueres Konzept» sei, da die «Bewertung bzw. Interpretation von Ergebnissen [...] immer einen subjektiven Vorgang [darstellt], der Ermessensspielraum beinhaltet und zu Diskussionen über Formulierungen im Evaluationsbericht Anlass geben kann, weil Evaluierende und Auftraggeber unterschiedliches Kontext- und Hintergrundwissen mitbringen und deshalb gegebenenfalls zu unterschiedlichen Interpretationen gelangen können». Ferner argumentierte der Evaluierende, dass es in solchen Situationen effektiv eine «Grauzone zwischen konstruktivem Dialog und Beeinflussung» gebe. Demnach beinhalteten Evaluationsprozesse «stets eine Komponente des Dialogs, des Aushandelns und des gegenseitigen Entgegenkommens», weshalb es eine «absolute Unabhängigkeit» im Sinne völliger Freiheit des Vorgehens, der Interpretation und der Berichterstattung» nicht gebe – zumindest nicht «im Kontext von Evaluation als Dienstleistung im Auftragsverhältnis».

Die Ergebnisse der offenen Rückmeldungen zeigten zudem, dass Beeinflussung oftmals subtil, d. h. implizit, geschieht (z. B. «Die Beeinflussungen sind viel subtiler, als im Fragebogen nachgefragt.»). Als Beispiele hierfür genannt wurden: langwieriges Herumfeilen an Fragebogen und Berichten, Verzögerungen oder Hinweise darauf, wie andere Evaluierende das gemacht hätten.

### 3 Ein heuristisches Modell der Beeinflussung

Die Rückmeldungen machen deutlich, dass – ebenso wie der Begriff der Unabhängigkeit (vgl. für eine begriffliche Klärung Widmer 2012) – der Begriff der Beeinflussung bislang unscharf definiert ist. Deshalb schlagen wir zu seiner konzeptionellen Erfassung ein neues Modell vor: das heuristische DRSB-Beeinflussungsmodell (*Distortion, Restraint, Support, Betterment: DRSB-heuristic model of influence*). Darin

wird Beeinflussung in die beiden Dimensionen «Entfaltungsgrad von Beeinflussung» (*Explicitness of influence*) und «Beeinflussungsintention» (*Direction of influence*) gegliedert. Der Entfaltungsgrad von Beeinflussung unterscheidet zwischen einer expliziten und einer impliziten Form von Beeinflussung, wobei die beiden Formen sich nach der Intensität der Beeinflussung voneinander abgrenzen. Ein impliziter Beeinflussungsgrad meint hierbei eine subtile Beeinflussung, wohingegen ein expliziter Grad eine offensichtliche Beeinflussung beschreibt. Die Beeinflussungsintention kann entweder positiv (konstruktiv) oder negativ (destruktiv) sein. Sie ist aus der Perspektive der Forschung zu verstehen, unterscheidet also danach, ob die Intention der wissenschaftlichen Forschung zuträglich oder abträglich ist. Abbildung 3 zeigt das heuristische DRSB-Modell, worin vier Beeinflussungstypen unterschieden werden: Verzerrung (*Distortion*), Zügelung (*Restraint*), Unterstützung (*Support*) und Verbesserung (*Betterment*).

Abbildung 3: Heuristisches DRSB-Modell

		Entfaltungsgrad der Beeinflussung ( <i>Explicitness of influence</i> )	
		Explizit ( <i>explicit</i> )	Implizit ( <i>implicit</i> )
Beeinflussungsintention ( <i>Direction of influence</i> )	Destruktiv ( <i>destructive</i> )	<b>Verzerrung (<i>Distortion</i>)</b>  Negative Beeinflussung	<b>Zügelung (<i>Restraint</i>)</b>  Negative Beeinflussung
	Konstruktiv ( <i>constructive</i> )	<b>Verbesserung (<i>Betterment</i>)</b>  Positive Beeinflussung	<b>Unterstützung (<i>Support</i>)</b>  Positive Beeinflussung

Hinsichtlich des heuristischen DRSB-Modells entsprechen einige der abgefragten Beeinflussungsformen eindeutig der explizit destruktiven Verzerrung. Hierunter fallen die Aufforderung des Auftraggebers zur positiveren oder negativeren Ergebnisdarstellung, die Aufforderung zu Änderungen, die zu inhaltlichen Verzerrungen führen würden, sowie die Mitteilung des Auftraggebers vor der Evaluation, welche Ergebnisse er oder sie erwartet. Die anderen abgefragten Beeinflussungsformen entsprechen der implizit destruktiven Zügelung.

Die von den Evaluierenden angegebenen Beeinflussungsformen lassen sich den vier Beeinflussungstypen des heuristischen DRSB-Modells zuordnen (Abbildung 4). Darin ist auch eine Auswahl von Beispielen für die beiden positiven Be-

einflussungsformen «Verbesserung» (*Betterment*) und «Unterstützung» (*Support*) aufgeführt, die zwar von den Evaluierenden genannt wurden, jedoch nicht primärer Gegenstand dieses Beitrags sind.

Abbildung 4: Zuordnung von erlebten Beeinflussungsformen zum DRSB-Modell

Entfaltungsgrad Beeinflussung ( <i>Explicitness of influence</i> )			
	Explizit ( <i>explicit</i> )	Implizit ( <i>implicit</i> )	
Beeinflussungsintention ( <i>Direction of influence</i> )	Destruktiv ( <i>destructive</i> )	<b>Verzerrung (<i>Distortion</i>)</b> <ul style="list-style-type: none"><li>• Aufforderung zur falschen Dateninterpretation</li><li>• Anpassung der Bewertung der Ergebnisse</li><li>• Anpassung der Darstellung der Ergebnisse, um einen bestimmten Schwellenwert zu erreichen</li><li>• Aufforderung, den qualitativen Teil vollständig zu streichen (methodische Mängel)</li><li>• Aufforderung, die Bewertungsreferenz im Nachhinein anzupassen</li><li>• Mitteilung, welche Politikempfehlungen der Schlussbericht beinhalten soll</li><li>• Beeinflussbare Skalen</li></ul>	<b>Zügelung (<i>Restraint</i>)</b> <ul style="list-style-type: none"><li>• Ablehnung oder Nicht-Veröffentlichung von Ergebnissen oder der gesamten Evaluation</li><li>• Vorgabe bzw. Ausschluss von Gesprächspartnern</li><li>• Langwieriges Herumfeilen an Fragebogen und Berichten</li><li>• Hinweise darauf, wie andere Evaluierende das gemacht hätten</li><li>• Festlegung gewisser Aspekte durch den Auftraggeber</li><li>• Ungeeignete methodische Vorgaben</li><li>• Schlecht formulierte Pflichtenhefte</li><li>• Honorar neu verhandeln</li><li>• Verlangen von <i>political correctness</i></li><li>• Suggestive Fragestellungen</li><li>• Zurückhalten wichtiger Informationen</li><li>• Absprechen der Qualität durch Auftraggeber</li><li>• Beeinflussung im Vorfeld, z. B. geschönte Darstellung des Gegenstands</li></ul>
	Konstruktiv ( <i>constructive</i> )	<b>Verbesserung (<i>Betterment</i>)</b> <ul style="list-style-type: none"><li>• Aufzeigen von Verbesserungspotenzial</li><li>• Austausch mit Auftraggeber über Ergebnisdarstellung zur Steigerung der Qualität einer Evaluation</li><li>• Präsentation von (Zwischen-)Ergebnissen, damit diese für das Zielpublikum verständlich sind, ohne die Resultate zu verfälschen</li></ul>	<b>Unterstützung (<i>Support</i>)</b> <ul style="list-style-type: none"><li>• Diskurs und Kompromissfindung</li><li>• Konstruktiver Dialog</li><li>• Darlegung der Sichtweise zu den Ergebnissen und zu deren Darstellung</li></ul>

Die Kategorisierung der offenen Rückmeldungen zeigt, dass insbesondere der destruktive implizite Typ von Beeinflussung, die Zügelung (*Restraint*), in der Evaluationspraxis häufig aufzutreten scheint. Dies entspricht den Ergebnissen der geschlossenen Fragen, wonach die von Evaluierenden am häufigsten erlebte Beeinflussungsform in der Aufforderung zur Umformulierung einzelner werbender Sätze bestand. Jedoch zeigen die offenen Antworten ein grösseres Spektrum an Ausprägungen der Zügelung, als durch die geschlossenen Fragen abgefragt wurde. Ein Evaluierender nannte beispielsweise die «Datenbeeinflussung» als problematische Form von Beeinflussung, worunter etwa Beeinflussung im Vorfeld gehöre. Ein anderer Evaluierender gab an, dass Beeinflussung über Methodik und Ergebnisse hinausreiche und bei der «Bewertung der Ergebnisse als Kern der Evaluation» ebenfalls eine grosse Rolle spiele.



Was die Beeinflussungsform der Zügelung betrifft, so wiesen mehrere Evaluierende darauf hin, dass Beeinflussung oftmals sehr subtil geschieht. Die Forderungen des Auftraggebers würden demnach weniger Forderungen im Sinne eines Verlangens darstellen, sondern vielmehr Wünsche: «Das Wort ‹verlangen› [ist] sehr restriktiv», da «in der Praxis [...] im Einzelfall ‹Wünsche› viel subtiler platziert [werden] denn als ‹Forderung›.» Ein anderer Evaluierender schrieb: «Es ist selten ein explizites Verlangen.» Weitere mehrfach aufgeführte Ausprägungen der Zügelung waren die Streichung einzelner Evaluationsteile oder die Ablehnung oder Nicht-Veröffentlichung der Evaluation. Hierbei wurde beispielsweise argumentiert, dass durch die Möglichkeit des Auftraggebers, die Veröffentlichung einer Evaluation zu verhindern, dieser einen kleineren Anreiz hat, die Evaluationsergebnisse zu beeinflussen (z. B. «Ein Aspekt der Beeinflussung [...] betrifft die Frage der Publikation des Evaluationsberichts. Gemäss den AGB des Bundes [...] kann die Publikation des Evaluationsberichts verhindert werden. [...] Damit ist die Versuchung des Auftraggebers kleiner, die Evaluationsergebnisse zu beeinflussen, da er deren Publikation verhindern kann»). Weitere Beispiele für die Beeinflussungsform der Zügelung waren die Vorgabe oder der Ausschluss von Gesprächspartnern oder das Hinauszögern der Evaluation durch langwieriges Herumfeilen an Fragebogen und Berichten.

Bei der Beeinflussungsform der Verzerrung (*Distortion*) nannten die Evaluierenden mehrfach die Aufforderung eines Stakeholders zur falschen Dateninterpretation oder zur Anpassung der Thematik- oder Ergebnisdarstellung. Weitere Beispiele für die Beeinflussungsform der Verzerrung waren die Aufforderung zur nachträglichen Anpassung der Bewertungsreferenz oder die vorgängige Angabe von gewünschten Ergebnissen in Form bestimmter Politikempfehlungen des Schlussberichts.

#### **4 Präventive Massnahmen**

Im Fragebogen wurden die Evaluierenden danach gefragt, ob es ihrer Ansicht nach Massnahmen gegeben hätte, die im Falle der Beeinflussung gegriffen hätten. Diese Frage beinhaltete neben einem geschlossenen Teil (Abbildung 5), in dem die Evaluierenden ihre Zustimmung zu möglichen präventiven Massnahmen angeben konnten, auch einen offenen Frageteil. Darin konnten weitere Massnahmen aufgeführt werden, die nach Ansicht der Evaluierenden geholfen hätten, den Beeinflussungsvorfall zu verhindern oder diesem entgegenzuwirken. Diese offene Frage nach präventiven Massnahmen ist auch deshalb relevant, da sich auf Basis der Antworten ein konkreter praktischer Nutzen für die Evaluationstätigkeit erarbeiten lässt.

*Abbildung 5: Zustimmung zu präventiven Massnahmen zur Verhinderung von Beeinflussung (N = 51)*



*Anmerkung:* Angaben in Prozent, Mehrfachnennungen möglich (in Anlehnung an Pleger/Sager 2016)

Die in den geschlossenen Fragen bereits hohe Zustimmung zur Schaffung eines gegenseitigen Verständnisses zwischen Auftraggeber und Auftragnehmer (51 % der Evaluierenden stimmten dieser Kategorie zu) wurde in den offenen Rückmeldungen mit Beispielen untermauert. Dadurch wird deutlich, wie wichtig dieses gegenseitige Verständnis aus Sicht der Evaluierenden für die Verhinderung von Beeinflussung ist. Die offenen Rückmeldungen zeigen ausserdem, dass viele Evaluierende der Auffassung sind, dass Auftraggeber oftmals gar nicht bewusst beeinflussen, sondern dies vielmehr aus mangelnder Kenntnis der Evaluationstät-

tigkeit oder Unwissenheit im Umgang mit Datengewinnung und Interpretation heraus geschieht. Mehrfach gaben die Evaluierenden an, dass eine Erhöhung des Evaluations-Knowhows auf Seiten der Auftraggeber dazu beitragen könnte, das Auftreten von Beeinflussung zu verringern oder gar zu verhindern. Diese Wahrnehmung lässt sich mit dem nachfolgenden Zitat eines Evaluierenden zusammenfassen:

Beeinflussung von Evaluationsvorhaben [...] entsteht oft durch den Umstand, dass die auftraggebenden Projektverantwortlichen wenig bis fast gar keine methodischen und fachspezifischen Kompetenzen haben. Die Beeinflussung ist somit nicht «absichtlich», sondern eine Konsequenz von fehlenden Kompetenzen.

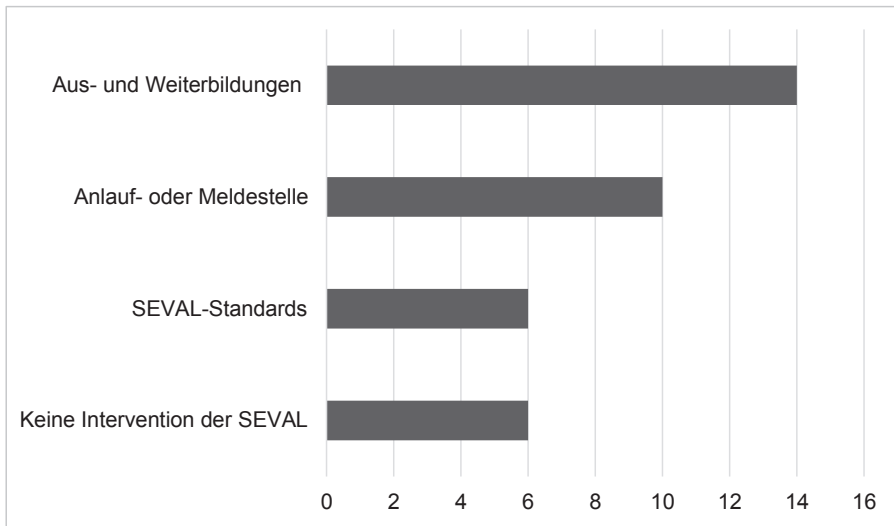
Neben indirekten Massnahmen in Form eines verbesserten gegenseitigen Verständnisses oder von besser ausgebildeten Stakeholdern schlugen die Evaluierenden in den offenen Fragen auch konkrete direkte Massnahmen vor. Ein Beispiel hierfür ist ein für Auftraggeber und Evaluierende obligatorisches Protokoll, worin beide Parteien unterschreiben, über die Evaluationsstandards und die Verpflichtung, diese einzuhalten, informiert zu sein, und zustimmen, sich an diese zu halten. Ein anderer konkreter Vorschlag eines Evaluierenden ist die Möglichkeit, die Auftragsformulierung zu revidieren.

Neben Vorschlägen zu direkten und indirekten präventiven Massnahmen waren drei Evaluierende der Auffassung, dass keine Massnahme gegriffen hätte, da ihrer Meinung nach nichts hätte unternommen werden können, was die Beeinflussung verhindert hätte.

## **5 Rolle der SEVAL und der Evaluationsstandards**

Die Ergebnisse der Studie über die Beeinflussung von Schweizer Evaluierenden zeigten sowohl ein ausgeprägtes Berufsethos der Evaluierenden als auch einen hohen Bekanntheitsgrad und Stellenwert der SEVAL-Standards (siehe hierfür Pleger/Sager 2016). In diesem Zusammenhang wurde im Fragebogen auch danach gefragt, welche Rolle die SEVAL im Falle einer Beeinflussung einnehmen sollte. Zu dieser Frage gab es eine grosse Anzahl Rückmeldungen, deren Auswertung nachfolgend vorgestellt wird. Der Grossteil der Rückmeldungen lässt sich in eine von vier Kategorien einordnen: Aus- und Weiterbildungen für Auftraggeber und Evaluierende, eine Anlauf- oder Meldestelle in Konfliktsituationen sowie die Promotion der SEVAL-Standards. Als vierte Kategorie lassen sich diejenigen Rückmeldungen zusammenfassen, die keine Intervention der SEVAL erwarteten. Abbildung 6 stellt die absolute Häufigkeit der von den Evaluierenden genannten Vorschläge und die Rolle der SEVAL im Zusammenhang mit Beeinflussung dar.

Abbildung 6: Absolute Häufigkeit vorgeschlagener Funktionen der SEVAL im Falle von Beeinflussung



Die von den Evaluierenden am häufigsten genannten Funktionen und Aufgaben der SEVAL im Zusammenhang mit Beeinflussung waren Aus- und Weiterbildungen, sowohl für Evaluierende als auch für Auftraggeber. Wie bereits bei den möglichen präventiven Massnahmen wurde von den Evaluierenden in den offenen Rückmeldungen mehrfach gefordert, dass Auftraggebern bessere Kenntnisse der Evaluationstätigkeit vermittelt und diese in der Begleitung von Evaluationen geschult werden müssten, da es «nach wie vor Stellen gebe [...], die keine Ahnung davon haben», wie eine Evaluation zu begleiten oder «was von einer Evaluation zu erwarten ist». Durch gezielte Schulungen oder Workshops sollten diese Weiterbildungen den Auftraggebern beispielsweise aufzeigen, «was von einer Evaluation zu erwarten ist», ihnen die «Rolle von Evaluierenden wissenschaftlicher Forschungs- und Evaluationsprinzipien» vermitteln oder deutlich machen, «dass sie keinen Einfluss nehmen dürfen». Auch Coachings für Evaluierende wurden aufgeführt, um diese etwa im Umgang mit Beeinflussungsversuchen zu schulen (z. B. «In der Ausbildung Kommunikation und Stakeholdermanagement/-Integration verstärken»). Mehrfach wurde hierbei von den Evaluierenden der Begriff der Professionalisierung verwendet, die von den Auftraggebern und Evaluierenden sowie in Zusammenhang mit der gesamten Evaluation gefördert werden müsste. Eine solche Professionalisierung würde dazu beitragen «Beeinflussungsversuche zu erkennen, ihnen zu widerstehen und [die] Unabhängigkeit zu wahren».

Neben Schulungen für Auftraggeber oder Evaluierende wurden von den Evaluierenden auch Veranstaltungen wie Workshops genannt, in deren Zentrum der Austausch stehen sollte, um das gegenseitige Verständnis zu fördern (z. B. «Organiser des ateliers d'échange sur ce thème, cours de sensibilisation pour les mandants»).

Hinsichtlich der SEVAL-Standards (Widmer et al. 2000) wurden mehrfach die Weiterentwicklung und die Promotion bei «allen an Evaluationen Beteiligten» aufgeführt. Als Adressaten wurden mehrmals die Auftraggeber genannt, bei welchen die SEVAL-Standards sowie die Rolle von Evaluierenden besser bekannt gemacht werden sollten. Ferner sollte die SEVAL die «allgemeine Professionalisierung der Evaluation» fördern und klare Standards zur Unabhängigkeit formulieren «auf [die] man sich im Falle einer Beeinflussung beziehen» könnte. Es wurde auch vorgeschlagen, die SEVAL-Standards verpflichtend jedem Vertrag anzufügen.

Ebenfalls mehrfach genannt wurde die Einrichtung einer Melde- oder Anlaufstelle bei der SEVAL, bei welcher sich Evaluierende im Falle von Beeinflussung beraten lassen oder der gravierende Fälle von Beeinflussung gemeldet werden könnten. Mehrere Evaluierende schlugen vor, dieser Anlaufstelle eine beratende Funktion zukommen zu lassen, wobei auch die Möglichkeit bestehen sollte, «Konfliktfälle zumindest anonymisiert [zu] erörtern». In diesem Zusammenhang wurde zudem vorgeschlagen, dass die SEVAL eine «neutrale Vermittlerrolle» einnehmen und gegebenenfalls als Mediatorin oder Ombudsstelle fungieren sollte (z. B. «Rôle de conseil, de médiateur sur demande de l'une des parties»; «Könnte sie sich als Schlichtungs- oder Beratungsstelle anbieten»; «Rôle de consultant peut-être»; «Le rôle de médiateur serait parfait»).

In einigen Rückmeldungen wurde der SEVAL zudem eine informierende Rolle zugesprochen, etwa die «Evaluation gegen aussen verstärkt als Profession [zu] vermitteln», den «Zweck und Nutzen von Evaluation in der Öffentlichkeit bekannter» zu machen, die «Instrumentalisierung von Evaluationen [zu] thematisieren und [zu] ächten» oder das «Spannungsfeld der Interpretation auf[zu]zeigen». Hierunter fallen auch die Rückmeldungen, die sich von der SEVAL eine «klare Positionierung» im Hinblick auf eine mögliche Beeinflussung wünschten, um mittels «Aufklärung» zur Prävention beizutragen.

Ein weiterer Vorschlag seitens der Evaluierenden dazu, welche Rolle die SEVAL bei Beeinflussungsvorfällen einnehmen könnte, war die Einführung eines «Rates der Weisen», bestehend aus unabhängigen Evaluierenden, bei denen man sich «Zweitmeinungen» einholen könnte.

In den offenen Rückmeldungen wurde jedoch auch argumentiert, dass die SEVAL keine Rolle im Kontext der Beeinflussung einnehmen könne oder solle, unter anderem weil sie nicht in der Lage sei, «eine Rolle in Konflikten rund um eine konkrete Evaluation einzunehmen», oder weil «mit den Standards bereits genug

Argumentarium zur Durchsetzung von «good practice» und Qualität» bestünde. Eine relativistische Sichtweise zur Rolle der SEVAL zeigt sich schliesslich in der folgenden Rückmeldung:

Beeinflussung ist menschlich, und das Widerstehen resp. Erliegen der Beeinflussung ebenfalls. Mit wachsender Erfahrung weiss ein guter Evaluierender meist bereits bei der Auftragsvergabe/Briefing, ob seitens des Auftraggebers Beeinflussung zu erwarten ist oder nicht, und kann entsprechend reagieren (ablehnen, Vorsorge treffen, entsprechendes meth. Verfahren wählen etc.). Allerdings wird wohl jeder Evaluierende im Verlaufe der Karriere in solche Situationen reinlaufen (aus mangelnder Erfahrung, aus Notwendigkeit einen Auftrag zu erhalten, aus Unachtsamkeit oder falscher Einschätzung etc.). Mit Beeinflussungsversuchen umzugehen lernen ist m. E. notwendiger Teil einer Evaluatorenkarriere.

## 6 Diskussion

Die hohe Anzahl offener Rückmeldungen von Schweizer Evaluierenden zum Thema Beeinflussung sowie zur Rolle der SEVAL in diesem Zusammenhang weist darauf hin, dass seitens der Evaluierenden zum einen kein Konsens über ein begriffliches Verständnis der Konzepte «Beeinflussung» und «Unabhängigkeit» vorhanden ist und zum anderen Bedarf an praktischen Weiterentwicklungen in diesem Bereich besteht.

Hinsichtlich der SEVAL-Standards könnte die explizite Aufnahme der Thematik der Unabhängigkeit dabei helfen, einen Konsens über den Begriff zu etablieren und dessen Bedeutung in der Praxis zu klären. Hierbei ist jedoch der Vieldeutigkeit der Begriffe «Unabhängigkeit» und «Beeinflussung» Rechnung zu tragen, da Beeinflussung in der Wahrnehmung und auch in der tatsächlichen Evaluationspraxis sowohl positiv als auch negativ und so für eine evidenzbasierte Evaluation förderlich, aber auch hinderlich sein kann. Nicht jede Einflussnahme eines Auftraggebers auf den Evaluierenden bedeutet eine Reduktion der wissenschaftlichen Unabhängigkeit einer Evaluation.

Die Ergebnisse zeigen, dass eine eindeutige Differenzierung selbst für die Evaluierenden schwierig ist. Diese Schwierigkeit lässt sich durch das nachfolgende Zitat eines Evaluierenden zusammenfassen:

Austausch mit den Auftraggebenden (auch über Ergebnisse) ist grundsätzlich wichtig (über Zielsetzungen, auch für die Einordnung von Ergebnissen) und hat viel Positives resp. ist für die Durchführung und letztlich auch für die Qualität einer Evaluation entscheidend. Dabei zwischen «negativer Beeinflussung» und «gehaltvollen, erwünschten Kommentaren/Inputs» zu unterscheiden, ist nicht immer einfach, vor allem wenn die Beeinflussung subtiler daherkommt.

Das Postulat nach einer unabhängigen Evaluation im Sinne einer Absenz von Beeinflussung als Massnahme zur Erhöhung der wissenschaftlichen Objektivität und Neutralität greift deshalb zu kurz. Vielmehr müssten für die Evaluations-

praxis Mechanismen entwickelt und etabliert werden, die eine Trennlinie zwischen einer negativen und glaubwürdigkeitsabträglichen Abhängigkeit in Folge der Einflussnahme und einer positiven und konstruktiven Beeinflussung zulassen. Konkrete Beispiele für solche Mechanismen wurden von den Evaluierenden selbst im Zusammenhang mit präventiven Massnahmen vorgeschlagen. Hierunter wurden oftmals auch Schulungen für die Auftraggeber oder andere Massnahmen, die sich speziell an Auftraggeber richten sollten, genannt. Evaluierende stehen in einer Austauschbeziehung mit ihren Auftraggebern. Diese gilt es von beiden Seiten aus bestmöglich vor Einflussnahme auf die wissenschaftsbasierte Evaluationspraxis selber und die daraus resultierenden Evaluationsergebnisse zu schützen. Deshalb muss sichergestellt werden, dass für eine moralische Verpflichtung hinsichtlich der Einhaltung fester Evaluationsstandards, wie sie bereits bei einem Grossteil der Auftragnehmerseite vorhanden ist, auch auf Auftraggeberseite gesorgt wird. Die Befragung von Auftraggebern war nicht Gegenstand der Studie, doch lassen sich daraus dennoch zwei wichtige Schlussfolgerungen hinsichtlich der Bedeutung von Auftraggebern beim Auftreten von Beeinflussung ableiten: Zum einen wurde die Relevanz der Auftraggeber in diesem Zusammenhang dadurch ersichtlich, dass die Auftraggeber von den Evaluierenden als die am häufigsten beeinflussenden Akteure wahrgenommen wurden (Pleger/Sager 2016). Zum anderen zeigten die Ergebnisse auch, dass Beeinflussung durch den Auftraggeber oftmals nicht bewusst geschieht. Bei den Rückmeldungen zu möglichen präventiven Massnahmen zeigte sich, dass aus Sicht vieler Evaluierender eine Beeinflussung durch den Auftraggeber aufgrund mangelnder Kenntnisse des Evaluationsprozesses heraus entsteht. Die Frage, ob sich «bei der Situation der Beeinflussung die Person darüber bewusst war, dass sie dazu anhielt, Ergebnisse falsch oder ungenau darzustellen», beantwortete lediglich ein Fünftel der Evaluierenden mit Ja. 43 Prozent gaben an, dass es sowohl Situationen gab, in denen sich die beeinflussende Person dessen bewusst war, als auch solche, in denen sie sich dessen nicht bewusst war. Rund ein Drittel beantwortete die Frage mit «weiss nicht» und 9 Prozent mit Nein (Pleger/Sager 2016). Diese Ergebnisse sind insofern frappierend, als die Beeinflussung demzufolge in etwa der Hälfte der Fälle überhaupt nicht beabsichtigt war. Bei diesen 50 Prozent könnten bereits präventive Massnahmen mittels Aufklärung und Information dazu beitragen, Beeinflussung zu verhindern.

Trotz oder gerade aufgrund dieser Schlussfolgerungen wird ebenfalls deutlich, dass eine zukünftige Befragung der Auftraggeberseite unbedingt erforderlich ist. Bislang nicht im DRSB-Modell enthalten ist die Beeinflussungsintention im Sinne einer bewussten oder unbewussten Beeinflussung. Durch die Einteilung in konstruktiv und destruktiv wird zwar die Beeinflussungsrichtung erfasst.

Diese stellt jedoch die Folgen der Beeinflussung dar und sagt nichts über die zugrunde liegende Intention des beeinflussenden Akteurs aus.

Eine weitere wichtige Schlussfolgerung aus den hier vorgestellten Ergebnissen betrifft den Kontext des Evaluationsprozesses, durch den Beeinflussung begünstigt oder verringert werden kann. Konkret ist damit der Auftrag, in dessen Rahmen eine Evaluation durchgeführt wird, gemeint. Stockmann et al. (2011) argumentieren, dass externe Evaluierende häufiger mit Beeinflussung konfrontiert sind, weil sie abhängiger von Folgeaufträgen sind und Auftraggeber diese Abhängigkeit antizipieren. Diese Annahme konnte in der Schweizer Studie quantitativ nicht bestätigt werden, was jedoch auch durch die Stichprobengrösse bedingt sein kann (Pleger/Sager 2016). So finden sich in den offenen Rückmeldungen durchaus Aussagen, welche die Annahme von Stockmann et al. (2011) stützen. Ein Evaluierender kommt beispielsweise zum Schluss, dass die «wirtschaftliche Abhängigkeit von Auftraggebern unter Umständen mehr beeinflussend ist als der Auftraggeber selbst.»

Die offenen Rückmeldungen zeigen auch, dass die SEVAL aus der Sicht vieler Evaluierender wichtig ist für die Entwicklung und Umsetzung präventiver Massnahmen zur Verhinderung von Beeinflussung. Demnach könnte die SEVAL diesbezüglich eine tragende Rolle einnehmen, etwa durch die Promotion der Standards oder durch deren Vermittlung sowohl an die Auftragnehmer- als auch an die Auftraggeberseite im Rahmen von Aus- und Weiterbildungen. Weiter könnte die SEVAL im Zusammenhang mit Schlichtungsprozessen als Meldestelle agieren, sofern etwa die SEVAL-Standards durch den Evaluierenden und seinen Auftraggeber als Vertragsbestandteil implementiert wurden. Auf die Ausgestaltung von Verträgen bei Evaluationsaufträgen kann die SEVAL keinen Einfluss nehmen.

*Lyn Pleger, MA PMP, Doktorandin, Kompetenzzentrum für Public Management, Universität Bern, Bern; E-Mail: Lyn.Pleger@kpm.unibe.ch*

*Fritz Sager, Prof. Dr., Professor für Politikwissenschaft, Kompetenzzentrum für Public Management, Universität Bern, Bern; E-Mail: Fritz.Sager@kpm.unibe.ch*



## Literaturverzeichnis

- Lincoln, Yvonna S. / Guba, Egon G., 1985, *Naturalistic Inquiry*, Thousand Oaks.
- Morris, Michael / Clark, Brittany, 2013, You want me to do WHAT? Evaluators and the pressure to misrepresent findings, *American Journal of Evaluation*, 34(1), S. 57–70.
- Pleger, Lyn / Sager, Fritz, 2016, „Don't tell me cause it hurts“ – Beeinflussung von Evaluierenden in der Schweiz, *Zeitschrift für Evaluation*, forthcoming.
- Prakke, Heleen / Wurster, Jahin, 1999, Gütekriterien für qualitative Forschung. *Pflege*, 12(3), S. 183–186.
- Raithel, Jürgen, 2006, *Quantitative Forschung*: Springer.
- Stockmann, Reinhard / Meyer, Wolfgang / Schenke, Hanna, 2011, Unabhängigkeit von Evaluationen, *Zeitschrift für Evaluation*, 10(1), S. 17–38.
- The LSE GV314 Group; Hayward, Rui Jin / Kay, Jonathan / Lee, Annabel / Page, Edward C / Patel, Nikesh / Payne, Hannah / Santos, Joana Pelicano / Perry, Lucy / Valyraki, Athina, 2014, Evaluation under contract: Government pressure and the production of policy research. *Public Administration* 92(1), S. 224–239.
- Widmer, Thomas, 2012, Unabhängigkeit in der Evaluation, *LeGes* 2012/2, S. 129–147
- Widmer, Thomas / Landert, Charles / Bachmann, Nicole, 2000, *Evaluations-Standards der Schweizerischen Evaluationsgesellschaft (SEVAL-Standards)*, Retrieved 24.8.2014.

## Résumé

*Peut-il arriver qu'un évaluateur soit confronté de la part du commanditaire de l'évaluation ou d'une autre partie prenante à une tentative d'influer sur son travail ? Récemment interrogés dans le cadre d'une enquête consacrée à cette question, un nombre appréciable d'évaluateurs actifs en Suisse ont affirmé s'être déjà trouvés dans cette situation. Mais que faut-il entendre au juste par « influence », terme généralement défini de manière peu précise, tout particulièrement lorsqu'il est question de l'indépendance des évaluations ? Telle est la question fondamentale à laquelle la présente contribution se propose de répondre en trois points. D'abord, elle procède à une analyse des réponses ouvertes faites par des évaluateurs suisses à la question de savoir ce qu'ils entendent par « influence ». Ensuite, elle présente le modèle DRSB de l'influence, qui tente au travers d'une approche heuristique de distinguer entre les différents types d'influence auxquels les évaluateurs ont été exposés. Enfin, elle dresse la liste des mesures qui aux yeux des évaluateurs seraient susceptibles de prévenir le risque d'influence et indique le rôle que le SEVAL pourrait jouer à cet égard. Précisons encore que l'analyse qualitative des réponses montre que l'influence s'exerce souvent de manière subtile, qu'il est difficile dans la pratique de faire le départ entre influence positive et influence négative, et que plusieurs évaluateurs suisses ont des vues très semblables quant à ce que devrait être l'action du SEVAL face à un cas d'influence.*